

EXPLICIT SOLUTIONS OF THE \mathfrak{a}_1 -TYPE LIE-SCHEFFERS SYSTEM AND A GENERAL RICCATI EQUATION

GABRIEL PIETRZKOWSKI

ABSTRACT. For a general differential system $\dot{x}(t) = \sum_{d=1}^3 u_d(t)X_d$, where X_d generates the simple Lie algebra of type \mathfrak{a}_1 , we compute the explicit solution in terms of iterated integrals of products of u_d 's. As a byproduct we obtain the solution of a general Riccati equation by infinite quadratures.

1. INTRODUCTION

Let M be a finite dimensional manifold. Let $X_a, X_b, X_c \in \Gamma(M)$ be smooth tangent vector fields on M , generating the \mathfrak{a}_1 -type Lie subalgebra, of the Lie algebra $\Gamma(M)$, such that

$$(1) \quad [X_a, X_b] = 2X_a, \quad [X_a, X_c] = -X_b, \quad [X_b, X_c] = 2X_c,$$

where $[\cdot, \cdot] : \Gamma(M) \otimes \Gamma(M) \rightarrow \Gamma(M)$ is a standard Lie bracket of tangent vector fields. Now for arbitrary measurable functions $u_a, u_b, u_c : [0, T] \rightarrow \mathbb{R}$, consider a differential equation on M defined as follows:

$$(2) \quad \begin{aligned} \dot{x}(t) &= u_c(t)X_c + u_b(t)X_b + u_a(t)X_a, \\ x(0) &= x_0 \in M. \end{aligned}$$

The goal is to write a general (local) solution for this system in terms of flows of $X_a, X_b, X_c \in \Gamma(M)$. This solution will depend explicitly on the functions u_a, u_b, u_c and their iterated integrals only.

The motivation to consider such a problem comes from many directions. For example, when u_a, u_b and u_c are constrained so as to define a control system, the solution described above gives rise to an explicit formula for the endpoint mapping of the control system [AS04].

Another example is when $M = \mathrm{SL}(2)$ is the special linear group, and X_d are left invariant vector fields. In this case the connections with Riccati equation are well known (see [Red56, Red57] and [CLR07a, Cd11]), as well as the subsequent applications to the study of Milne-Pinney equations (see [CdL09]), Ermakov systems (see [CdLR08]), and harmonic oscillators (see [CdLR10]). In particular our approach gives rise to the explicit solution of a general Riccati equation (see Theorem 2).

The advantage of our approach is that it can be generalized to systems of any simple-lie-algebra-type, in particular it will be very interesting to write out similar solutions for rank-two simple Lie algebras.

In order to solve the stated problem, we use the theorem of Kawski and Sussmann [KS97]. The origins of their solution, come from the works of KT Chen [Che54, Che57, Che68] on iterated integration of paths, also called algebraic paths, and later application of Chen's results to control systems by Fliess [Fli81]. An important step was given by Sussmann [Sus86], and in strictly algebraic fashion by Melançon and Reutenauer [MR89], where he expressed his solution of non-linear control-affine system, in terms of a Lyndon

2010 *Mathematics Subject Classification.* 17B80, 34A05, 34A26.

Key words and phrases. free Lie algebra, shuffle product, special linear algebra, Riccati equation, Lie-Sheffers system.

basis. This was then generalized in Reutenauer's book [Reu93] (who actually claims that the original idea comes from Schützenberger and Malançon), and rewritten in a control-theoretical setting by Kawski and Sussmann [KS97].

The article is organized as follows. In section 2 we formulate the results; in section 3 we collect important preliminary facts from the theory of free Lie algebras and its connection with integration of differential systems; in section 4 we introduce a Hall set/basis adjusted to our problem, which allows us to prove our results in section 5. Finally, we conclude the paper in section 6.

2. RESULTS

Before we state the main theorem, we need a few definitions. Let $A = \{a, b, c\}$ be a set of letters. By A_n^* denote the set of words composed from n letters, i.e.

$$A_n^* = \{a_1 a_2 \cdots a_n \mid a_1, \dots, a_n \in A\},$$

and by $A^* = \bigcup_{n \in \mathbb{N} \cup \{0\}} A_n^*$ the set of all words. In particular $A_0^* = \{1\}$ contains only one – empty – word. It is important to note that A^* , with an associative, noncommutative product (called the concatenation product) given by juxtaposition $(v, w) \mapsto vw$, and the neutral element 1, is a monoid. We denote by $|w|$ the length of $w \in A^*$, i.e. $|w| = n$ for $w \in A_n^*$, and by $|w|_d$ the number of occurrences of the letter d in the word $w \in A^*$.

The concatenation product gives rise to the \mathbb{R} -algebras $\mathbb{R}\langle A \rangle$ of noncommutative polynomials on A , and $\mathbb{R}\langle\langle A \rangle\rangle$ of non-commutative power series on A . We denote by $\sqcup : \mathbb{R}\langle\langle A \rangle\rangle \otimes \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$ the standard bilinear shuffle product (see section 3 for the definition), and denote by $\mathbb{R}_{\sqcup}\langle\langle A \rangle\rangle$ the commutative \mathbb{R} -algebra of power series on A with the product \sqcup (we use this notation only in case the shuffle algebra structure is crucial). We define the shuffle exponential $\exp_{\sqcup} : \mathbb{R}\langle\langle A \rangle\rangle_0 \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$ by

$$\exp_{\sqcup}(P) := 1 + \sum_{k=1}^{\infty} \frac{P^{\sqcup k}}{k!},$$

where the "0" subscript in $\mathbb{R}\langle\langle A \rangle\rangle_0$ means that we take P with zero constant term.

For fixed measurable controls $u_a, u_b, u_c : [0, t] \rightarrow \mathbb{R}$, we define the linear homomorphism $\Upsilon^t : \mathbb{R}_{\sqcup}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}$ given by

$$(3) \quad A^* \ni w = a_1 \cdots a_n \mapsto \Upsilon^t(w) := \int_0^t u_{a_n}(t_n) \int_0^{t_n} \cdots \int_0^{t_2} u_{a_1}(t_1) dt_1 \cdots dt_{n-1} dt_n.$$

It is easy to check that Υ^t is in fact an associative algebra homomorphism, i.e. $\Upsilon^t(v \sqcup w) = \Upsilon^t(v) \Upsilon^t(w)$ (see [Che68, Reu93, Kaw02]).

Finally denote by $\exp(tX) : M \rightarrow M$ the flow of a tangent vector field $X \in \Gamma(M)$ at time $t \geq 0$.

Theorem 1. *Let $X_a, X_b, X_c \in \Gamma(M)$ be smooth tangent vector fields satisfying (1). Then (locally) the solution $x : [0, T] \rightarrow M$ of the differential equation (2) is of the form*

$$(4) \quad x(t) = \exp(\Xi_c(t)X_c) \exp(\Xi_b(t)X_b) \exp(\Xi_a(t)X_a)(x_0).$$

Here, $\Xi_a, \Xi_b, \Xi_c : [0, T] \rightarrow \mathbb{R}$ are given by $\Xi_d(t) := \Upsilon^t(S^d)$ (for $d = a, b, c$), where

$$S^a = a \exp_{\sqcup}(2S_{a_1}), \quad S^b = S_{a_1}, \quad S^c = \exp_{\sqcup}(2S_{a_1})c,$$

and $S_{a_1} \in \mathbb{R}\langle\langle A \rangle\rangle$ is the unique solution of the algebraic equation

$$S_{a_1} = b - a \exp_{\sqcup}(2S_{a_1})c.$$

In particular, we have

$$b - aS^c = S^b = b - S^a c.$$

Remark 2.1. The solution in the form (4) exists locally by virtue of the Wei and Norman theorem [WN64]. It is not our goal to investigate the issue of the global solution in this paper (in the mentioned article Wei and Norman give example of a local but non-global solution in $\mathfrak{so}(3)$ case), rather, we focus on the formal one.

Remark 2.2. Since $\Upsilon^t : \mathbb{R}_{\sqcup} \langle \langle A \rangle \rangle \rightarrow \mathbb{R}$ is an algebra homomorphism, it is easy to check that

$$\Xi_c(t) = \int_0^t u_c(t_1) \cdot e^{2\Upsilon^{t_1}(S_{\mathfrak{a}_1})} dt_1.$$

As a byproduct of the proof of the above theorem we get that Ξ_a is the solution of a Riccati equation. We state it in the following theorem.

Theorem 2. *For fixed measurable functions $u_a, u_b, u_c : [0, T] \rightarrow \mathbb{R}$ the function $\Xi_a : [0, T] \rightarrow \mathbb{R}$, defined in Theorem 1 by $\Xi_a(t) = \Upsilon^t(a \exp_{\sqcup}(2S_{\mathfrak{a}_1}))$, is (locally) the solution of the Riccati equation:*

$$\begin{aligned} \dot{\Xi}_a(t) &= u_a(t) + 2u_b(t) \cdot \Xi_a(t) - u_c(t) \cdot \Xi_a^2(t) \\ \Xi_a(0) &= 0. \end{aligned}$$

As is well known, a specific solution of a Riccati equation gives rise to all its other solutions (the appropriate formula can be found for example in [CLR07b]). An easy calculation gives the following corollary.

Corollary 2.3. *For fixed measurable functions $u_a, u_b, u_c : [0, T] \rightarrow \mathbb{R}$, the function $y : [0, T] \rightarrow \mathbb{R}$, where*

$$y(t) = \Upsilon^t(a \exp_{\sqcup}(2S_{\mathfrak{a}_1})) + \frac{y_0 \cdot e^{2\Upsilon^t(S_{\mathfrak{a}_1})}}{1 + y_0 \cdot \int_0^t u_c(t_1) \cdot e^{2\Upsilon^{t_1}(S_{\mathfrak{a}_1})} dt_1},$$

is (locally) the solution of the Riccati equation:

$$\begin{aligned} \dot{y}(t) &= u_a(t) + 2u_b(t) \cdot y(t) - u_c(t) \cdot y^2(t) \\ y(0) &= y_0. \end{aligned}$$

The series $S_{\mathfrak{a}_1} \in \mathbb{R} \langle \langle A \rangle \rangle$ can be computed explicitly. In order to do this, we recursively define two series of ordered sets of words $W_n^b, W_n^c \subset A_n^*$ ($n = 1, 2, \dots$) as follows (in each set we denote by \leq the according total ordering):

- (a) $W_1^b := \{b\}$ and for $n \geq 2$, $W_n^b = aW_{n-1}^c := \{av \mid v \in W_{n-1}^c\}$ with ordering taken from W_{n-1}^c that is $av < aw$ iff $v < w$.
- (b) $W_1^c := \{c\}$ and for $n \geq 2$,
 - (i) $W_n^c := \bigcup_{i=1}^k W_i^b W_{n-i}^c := \bigcup_{i=1}^k \{vw \mid v \in W_i^b, w \in W_{n-i}^c\}$ if $n = 2k$ is even, with "the lexicographical ordering" that is $vw < v'w'$ with $v, v' \in W_i^b$ and $w, w' \in W_{n-i}^c$ iff $v < v'$ or $(v = v' \text{ and } w < w')$;
 - (ii) $W_n^c := \bigcup_{i=1}^k W_i^b W_{n-i}^c \cup \Delta^k$ if $n = 2k + 1$ is odd, where

$$\Delta^k := \{avw \mid v, w \in W_k^c, v \geq w\}.$$

The ordering is like in the case (i) with additional relations that $v < w$ for all $v \in \Delta^k$, $w \in \bigcup_{i=1}^k W_i^b W_{n-i}^c$, and $avw < av'w'$ with $v, v', w, w' \in W_k^c$ iff $v < v'$ or $(v = v' \text{ and } w < w')$.

For $d = b, c$, define the set $W^d := \bigcup_{n=1}^{\infty} W_n^d$. In each of these sets we introduce the ordering originating from the orderings \leq on W_n^d with the additional relation that for $v \in W_p^d, w \in W_q^d$ with $p > q$ we put $v \leq w$.

For all words in $W := W^b \cup W^c \subset A^*$, we recursively define a mapping $s : W \rightarrow \mathbb{R} \langle A \rangle$ by

- (a) $s_b := b$ and
 $s_{av} := as_v$ for $av \in W_n^a$, where $v \in W_{n-1}^c$.
 (b) $s_c := c$,
 $s_{bw} := bs_w$ for $bw \in W_n^c$, where $w \in W_{n-1}^c$, and
 $s_{avw} = as_v \sqcup s_w$ for $avw \in W_n^c$, where $v, w \in W_{n-1}^c$.

The product \sqcup is a slight modification of \sqcup , in particular, for $v_1 = \dots = v_{i_1} > \dots > w_1 = \dots = w_{i_k}$, where $i_1, \dots, i_k \in \mathbb{N}$, we define

$$s_{v_1} \sqcup \dots \sqcup s_{w_{i_k}} := \frac{1}{i_1! \dots i_k!} s_{v_1}^{\sqcup i_1} \sqcup \dots \sqcup s_{w_1}^{\sqcup i_k},$$

where $P^{\sqcup 1} := P$ and $P^{\sqcup i} := P^{\sqcup i-1} \sqcup P$, $i = 2, 3, \dots$, for all $P \in \mathbb{R}\langle\langle A \rangle\rangle$.

Note that in case (b), the definition is enough, since for $vw \in W^c$, where $v \in W^b$ and $w \in W^c$, we have $v = av'$ with $v' \in W^b$.

Finally, we define the main series $S \in \mathbb{R}\langle\langle A \rangle\rangle$ by

$$(5) \quad S := \sum_{w \in W^b} \alpha_b(w) s_w,$$

where $\alpha_b(b) := 1$ and

$$(6) \quad \alpha_b(aw') := -(-2)^{|w'|_a} \cdot 2^{|w'|_b}$$

for $aw' \in W^b$.

Proposition 2.4. *The series $S \in \mathbb{R}\langle\langle A \rangle\rangle$ defined in (5), is the unique solution of the algebraic equation*

$$S = b - a \exp_{\sqcup}(2S) c.$$

In other words $S_{a1} = S$.

The expansion of the main series up to words of length less than 6, is

$$S = b - ac - 2abc - 4abbc + 2aacc - 8abbbc + 4a(acb + 2abc + bac)c + \dots$$

This expansion suggests a certain symmetry of the series S . Namely, define an algebra antihomomorphism $(\cdot)^* : \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$, $(a_1 \dots a_n)^* = a_n \dots a_1$ for $a_1 \dots a_n \in A^*$, and an algebra homomorphism $F : \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$, such that $F(a) = c$, $F(b) = b$, and $F(c) = a$. Denote by $F^* : \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}\langle\langle A \rangle\rangle$ the composition of $(\cdot)^*$ with F .

Proposition 2.5. *Let $S^a, S^b, S^c \in \mathbb{R}\langle\langle A \rangle\rangle$ be the series defined in Theorem 1. Then*

$$F^*(S^b) = S^b, \quad F^*(S^a) = S^c, \quad F^*(S^c) = S^a.$$

Finally, we compute an expression for $\exp_{\sqcup}(2S)$.

Proposition 2.6. *For the main series $S \in \mathbb{R}\langle\langle A \rangle\rangle$, it follows that*

$$\exp_{\sqcup}(2S) = 1 + \sum_{i=1}^{\infty} 2^i \sum_{\substack{w_1 \geq \dots \geq w_i \\ w_i \in W^b}} \alpha_b(w_1) \dots \alpha_b(w_i) s_{w_1} \sqcup \dots \sqcup s_{w_i}.$$

3. PRELIMINARY RESULTS

The proofs of Theorems 1 and 2, are based on the results of Kawski and Sussmann [KS97, section 5] which generalize Sussmann's work [Sus86]. Actually, a similar solution stated in a strictly algebraic fashion, can be found in [Reu93, Corollary 5.6] (the algebraic analog of the earlier Sussmann's result was proved in [MR89]). In order to state these results, we introduce necessary notions from the theory of free Lie algebras (we follow [Reu93] where the proofs and details can be found).

Assume $'A$ is a certain set which we call the alphabet. As in section 2, we denote by $'A_n^*$, $'A^*$, $\mathbb{R}\langle'A\rangle$, and $\mathbb{R}\langle\langle'A\rangle\rangle$, the set of words of length n , the set of all words, the \mathbb{R} -algebras of non-commutative polynomials, and series in the letters $'A$, respectively. Since $\mathbb{R}\langle'A\rangle \subset \mathbb{R}\langle\langle'A\rangle\rangle$, and $\mathbb{R}\langle\langle'A\rangle\rangle$ is the algebraic closure of $\mathbb{R}\langle'A\rangle$, we define all objects in the larger algebra. In particular, the product of two series $P = \sum_{w \in 'A^*} (P|w)w$ and $Q = \sum_{w \in 'A^*} (Q|w)w$ is defined by $PQ := \sum_{w \in 'A^*} (PQ|w)w$ with $(PQ|w) := \sum_{uv=w} (P|u)(Q|v)$. In $\mathbb{R}\langle\langle'A\rangle\rangle$ we define the standard Lie bracket $[\cdot, \cdot] : \mathbb{R}\langle\langle'A\rangle\rangle \otimes \mathbb{R}\langle\langle'A\rangle\rangle \rightarrow \mathbb{R}\langle\langle'A\rangle\rangle$ given by $[P, Q] := PQ - QP$ for $P, Q \in \mathbb{R}\langle\langle'A\rangle\rangle$. We denote by $\text{Lie}('A)$ the smallest \mathbb{R} -submodule of $\mathbb{R}\langle\langle'A\rangle\rangle$ which contains $'A$, and is closed under the Lie bracket. As is well known, $\text{Lie}('A)$ is the algebraic closure of the free Lie algebra generated by $'A$.

On $\mathbb{R}\langle\langle'A\rangle\rangle$ we also consider the bilinear shuffle product $\sqcup : \mathbb{R}\langle\langle'A\rangle\rangle \otimes \mathbb{R}\langle\langle'A\rangle\rangle \rightarrow \mathbb{R}\langle\langle'A\rangle\rangle$ defined recursively for words by putting $1 \sqcup w = w \sqcup 1 = w$ for any $w \in 'A^*$, and

$$(7) \quad (w_1 a_1) \sqcup (w_2 a_2) = (w_1 \sqcup (w_2 a_2)) a_1 + ((w_1 a_1) \sqcup w_2) a_2$$

for all $a_1, a_2 \in 'A$ and $w_1, w_2 \in 'A^*$. It is easy to check, that (7) is equivalent to

$$(8) \quad (a_1 w_1) \sqcup (a_2 w_2) = a_1 (w_1 \sqcup (a_2 w_2)) + a_2 ((a_1 w_1) \sqcup w_2)$$

for all $a_1, a_2 \in 'A$ and $w_1, w_2 \in 'A^*$.

Let $M('A)$ be the set of binary, complete, planar, rooted trees with leaves labelled by $'A$. Each such tree can be naturally identified with the unique expression in the set $E('A)$ defined by the following two conditions: (i) if $a \in 'A$, then $a \in E('A)$, and (ii) if $t, t' \in E('A)$, then $(t, t') \in E('A)$. In the sequel we will not distinguish between these sets, i.e. we assume $M('A) = E('A)$. Define the mapping $f(\cdot) : M('A) \rightarrow 'A^*$, which assigns to a tree $t \in M('A)$ the word given by dropping all brackets in it, i.e., $f(a) = a$ for all $a \in 'A$, and $f((t, t')) = f(t)f(t')$ for all $t, t' \in M('A)$. The word $f(t)$ is called the foliage of $t \in M('A)$. Define also the mapping $P : M('A) \rightarrow \text{Lie}('A)$, which changes the rounded brackets into the Lie brackets, i.e., $P_a = a$ for all $a \in 'A$, and $P_{(t, t')} := [P_t, P_{t'}]$ for all $t, t' \in M('A)$. We will generalize this definition in the sequel. A Hall set $'H$ on the letters $'A$ (which should also be called a Shirshov set [Shi62] and a Viennot set [Vie78]), is a subset of $M('A)$ totally ordered by \leq and satisfying:

- (I) $'A \subset 'H$;
- (II) if $h = (h', h'') \in 'H \setminus 'A$, then $h'' \in 'H$ and $h < h''$;
- (III) for all $h = (h', h'') \in M('A) \setminus 'A$ we have $h \in 'H$ iff
 - $h', h'' \in 'H$ and $h' < h''$, and
 - $h' \in 'A$ or $h' = (x, y)$ such that $y \geq h''$.

Fix a Hall set $'H$ on the letters $'A$ totally ordered by \leq . Each Hall tree $h \in 'H$ corresponds to a word $f(h) \in 'A^*$ called a Hall word. Denote by $'W$, the set of Hall words with ordering \leq inherited from the ordering on $'H$ in the natural way. It is a nontrivial fact that each word $w \in 'A^*$, is the unique concatenation of a unique non-increasing series of Hall words, that is, $w = h_1 \cdots h_k$ for some unique $k \in \mathbb{N}$, and $h_i \in 'W$ such that $h_1 \geq \cdots \geq h_k$ (in the sequel we will use letter $'h$ to describe both Hall words and Hall trees). Let $P : 'A^* \rightarrow \mathbb{R}\langle'A\rangle$ be the mapping defined by

- (i) $P_1 := 1$;
- (ii) $P_a := a$ for $a \in 'A$;
- (iii) $P_h := P_t \in \text{Lie}('A)$ for $h \in 'W$ such that $h = f(t)$, $t \in 'H \subset M('A)$;
- (iv) $P_w := P_{h_1} \cdots P_{h_k} \in \mathbb{R}\langle'A\rangle$ for $w = h_1 \cdots h_k$, where $k \in \mathbb{N}$ and $h_i \in 'W$ such that $h_1 \geq \cdots \geq h_k$.

The set $'B := \{P_h \in \text{Lie}('A) \mid h \in 'W\}$ is the Hall basis of $\text{Lie}('A)$ corresponding to the Hall set $'H$. By the Poincaré-Birkhoff-Witt theorem, the set of ordered products $P_{h_1} \cdots P_{h_k}$, where $h_1 \geq \cdots \geq h_k$ are Hall words, creates a basis for the enveloping algebra of $\text{Lie}('A)$, which in the free case is isomorphic to $\mathbb{R}\langle\langle'A\rangle\rangle$. Therefore $\{P_w \mid w \in 'A^*\}$ is a basis in

$\mathbb{R}\langle\langle A \rangle\rangle$ (but in fact each $P_w \in \mathbb{R}\langle A \rangle$, so it is also a basis in $\mathbb{R}\langle A \rangle$). For our purpose it is crucial to consider the dual basis $\{\xi_w \mid w \in A^*\}$ of the algebra $\mathbb{R}\langle\langle A \rangle\rangle$, defined as always by

$$v = \sum_{w \in A^*} (\xi_w | v) P_w$$

for any $v \in A^*$. In this context we have the following proposition.

Proposition 3.1 ([Reu93, Theorem 5.3]). (i) $\xi_1 = 1$;
(ii) If $h = av \in {}^A W$ is a Hall word, where $a \in A, v \in A^*$, then $\xi_h = a\xi_v$;
(iii) If $w = h_1^{i_1} \cdots h_k^{i_k} \in A^*$ is any word, where $h_1 > \cdots > h_k$ are Hall words and $i_1, \dots, i_k \in \mathbb{N}$, then

$$\xi_w = \frac{1}{i_1! \cdots i_k!} \xi_{h_1}^{\sqcup i_1} \sqcup \cdots \sqcup \xi_{h_k}^{\sqcup i_k}$$

(recall that $P^{\sqcup 1} = P$ and $P^{\sqcup i} = P^{\sqcup i-1} \sqcup P$, $i = 2, 3, \dots$, for all $P \in \mathbb{R}\langle A \rangle$).

In what follows we will sometimes use a natural notation $\xi_h := \xi_{f(h)}$ for $h \in H$.

In [KS97], Kawski and Sussmann consider the so called universal control system evolving in the algebra $\mathbb{R}\langle\langle A \rangle\rangle$, and the solution is known as the Chen-Fliess series. Their main result states that the Chen-Fliess series is equal to the infinite product of group-like elements in $\mathbb{R}\langle\langle A \rangle\rangle$ (i.e., elements $\exp(S)$ where $S \in \text{Lie}(A)$), parametrized by any Hall set on the letters A . For our purposes it is important that such a universal control system can be utilized in examining a differential affine control system:

$$(9) \quad \begin{aligned} \dot{y}(t) &= \sum_{a \in A} u_a(t) Y_a, \\ y(0) &= y_0 \in M, \end{aligned}$$

where M is an arbitrary finite dimensional smooth manifold, $u_a : [0, T] \rightarrow \mathbb{R}$ are measurable controls, and $Y_i \in \Gamma(M)$ are fixed smooth tangent vector fields. We state the Kawski-Sussmann theorem for this control system. Namely, if we take any Hall set H , then for fixed u_i 's (locally) the solution to (9) is

$$(10) \quad y(t) = \overleftarrow{\prod}_{h \in H} \exp(\Phi_h(t) \cdot Y_h)(y_0).$$

The symbol $\overleftarrow{\prod}_{h \in H}$ denotes the decreasing product with respect to the ordering \leq in H , i.e.

$$\overleftarrow{\prod}_{h \in H} F_h := \cdots F_h F_{h'} \cdots,$$

where $\cdots > h > h' > \cdots$. The tangent vector fields Y_h , are defined by the relation of the Hall set with the Hall basis in $\text{Lie}(A)$, and the universal property of the free Lie algebra $\text{Lie}(A)$, that is $Y : H \rightarrow \Gamma(M)$ is the composition of the mapping $H \ni h \mapsto P_h \in \text{Lie}(A)$ with the unique Lie-algebra homomorphism generated by $\text{Lie}(A) \supset A \ni a \mapsto Y_a \in \Gamma(M)$.

In order to define $\Phi_h(t)$, we introduce the Zinbiel algebra $\Pi_{[0,t]}(A)$ of iterated integrals of controls $u_a(t), a \in A$, that is the \mathbb{R} -algebra generated by $U_a(t) := \int_0^t u_a(t_1) dt_1$ with Zinbiel product $*$: $\Pi_{[0,t]}(A) \otimes \Pi_{[0,t]}(A) \rightarrow \Pi_{[0,t]}(A)$ given by

$$(f_1 * f_2)(t) := \int_0^t f_1(t_1) f_2'(t_1) dt_1.$$

The algebra is Zinbiel since $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3 + (f_2 * f_1) * f_3$ for all $f_1, f_2, f_3 \in \Pi_{[0,t]}(A)$, which is easy to check. In particular,

$$((\cdots (U_{a_1} * U_{a_2}) * \cdots) * U_{a_n})(t) = \int_0^t u_{a_n}(t_n) \int_0^{t_n} \cdots \int_0^{t_2} u_{a_1}(t_1) dt_1 \cdots dt_{n-1} dt_n.$$

Therefore, we define the linear mapping $\Upsilon_f^t : \mathbb{R}\langle\langle A \rangle\rangle \rightarrow \Pi_{[0,t]}(A)$ by putting $\Upsilon_f^t(1) = 1$ and

$$(11) \quad \begin{aligned} \Upsilon_f^t(a_1 \cdots a_n) &:= \Upsilon_f^t(a_1 \cdots a_{n-1}) * U_{a_n} \\ &= ((\cdots (U_{a_1} * U_{a_2}) * \cdots) * U_{a_n})(t) \end{aligned}$$

for any word $a_1 \cdots a_n \in A^*$. For fixed integrable controls u_a , we define $\Upsilon^t : \mathbb{R}_{\sqcup}\langle\langle A \rangle\rangle \rightarrow \mathbb{R}$, $S \mapsto \Upsilon^t(S)$ as the evaluation of $\Upsilon_f^t(S) \in \Pi_{[0,t]}(A)$ on these controls. This definition coincides with (3). Since ξ_h are expressed in terms of the shuffle product, it is important to note that [Che68, Reu93, Kaw02]

$$(12) \quad \Upsilon^t(v \sqcup w) = \Upsilon^t(v) \cdot \Upsilon^t(w),$$

where \cdot is the ordinary multiplication in \mathbb{R} .

Now the remaining definition in (10) is

$$(13) \quad \Phi_h(t) := \Upsilon^t(\xi_{f(h)}).$$

4. \mathfrak{a}_1 -TYPE HALL SET

In this section we begin to prove the theorems stated in section 2. As we pointed out in the previous section, we are going to use formula (10) of the formal solution of (9). In order to do that, we construct a specific Hall set $H^{\mathfrak{a}_1}$ on the set of letters $A = \{a, b, c\}$ ordered by \leq , adapted to Lie algebra of the type \mathfrak{a}_1 . Let $I(\mathfrak{a}_1)$ be the smallest Lie ideal of $\text{Lie}(A)$ generated by elements of the form $[a, b] - 2a$, $[a, c] + b$, $[b, c] - 2c$. Then $\mathfrak{A}_1 := \text{Lie}(A)/I(\mathfrak{a}_1)$ is a Lie algebra of the type \mathfrak{a}_1 . The Hall set $H^{\mathfrak{a}_1}$ is going to satisfy the following conditions:

- (I) $H^{\mathfrak{a}_1} = H^o \cup H^a \cup H^b \cup H^c$; if $h \in H^d$ ($d = a, b, c$), then
- (14) $P_h = \gamma_d(h) \cdot d \bmod I(\mathfrak{a}_1)$
- with certain functions $\gamma_d : H^d \rightarrow \mathbb{Z}$, and if $h \in H^o$, then $P_h = 0 \bmod I(\mathfrak{a}_1)$;
- (II) $H^d = \bigcup_{n \in \mathbb{N}} H_n^d$ for $d = o, a, b, c$, and $h \in H_n^d$ iff $h \in H^d$ and $|h| = n$;
- (III) each set H_n^d is finite.

Assuming we define H_n^d with total orderings \leq of the type (cardinality of $H_n^d \in \mathbb{N}$, we define the total ordering \leq in $H^{\mathfrak{a}_1}$ of the type $\omega \cdot 4$ (i.e., the ordering isomorphic to the lexicographical ordering on $\{c, b, a, o\} \times \mathbb{N}$) by adding the following intuitive relations:

- (A) for $h_1 \in H_{n_1}^d, h_2 \in H_{n_2}^d$ with $n_1 \neq n_2$ we have $h_1 < h_2$ iff $n_1 > n_2$;
- (B) for $h_o \in H^o, h_a \in H^a, h_b \in H^b, h_c \in H^c$ we have $h_o < h_a < h_b < h_c$.

Let us comment on these assumptions. The reason for distinguishing such a Hall set is that although we consider the free Lie algebra, to write out the solution of the differential equation, we actually assume certain relations on the vector fields involved. That is why we define the ideal $I(\mathfrak{a}_1)$, and the quotient algebra \mathfrak{A}_1 . Namely, if we extend the mapping $A \rightarrow \Gamma(M)$, $d \mapsto X_d$ for $d = a, b, c$, to the Lie algebra homomorphism $X : \text{Lie}(A) \rightarrow \Gamma(M)$, then $I(\mathfrak{a}_1)$ is in the kernel of X , so we can also consider the quotient algebra homomorphism $X : \mathfrak{A}_1 \rightarrow \Gamma(M)$. The point is that, by (I), $X_h = \gamma_d(h)X_d$ for each $h \in H^d$ if $d = a, b, c$, and $X_h = 0$ for each $h \in H^o$. In particular, we do not need to care about elements or the ordering of H^o .

Let us construct the sets H_n^d . By the definition, each Hall set contains all letters, so if (I) and (II) (for $n=1$) are to be satisfied, we define $H_1^d := \{d\}$ for $d = a, b, c$ and $H_1^o := \emptyset$. By (B) we have $a < b < c$. In order to satisfy the definition of a Hall set for $n = 2$, we must allocate $(a, b), (a, c), (b, c)$. To fulfill (I) we put

$$H_2^a := \{(a, b)\}, \quad H_2^b := \{(a, c)\}, \quad H_2^c := \{(b, c)\}, \quad H_2^o := \emptyset.$$

By (B), the ordering is $(a, b) < a < (a, c) < b < (b, c) < c$. We pass to $n = 3$. We need to add $((a, b), a), ((a, b), b), (a, (a, c)), (a, (b, c)), ((a, c), b), ((a, c), c), (b, (b, c)), ((b, c), c)$. Now it

is easy to compute Lie polynomials connected with these elements modulo $I(\mathfrak{a}_1)$, and to satisfy (I) we must put

$$\begin{aligned} H_3^a &:= \{(a, (a, c)), ((a, b), b)\}, & H_3^b &:= \{(a(bc))\}, \\ H_3^c &:= \{(b, (b, c)), ((a, c), c)\}, & H_3^o &:= \{((a, b), a), ((a, c), b), ((b, c), c)\}. \end{aligned}$$

We order these sets so that $(a, (a, c)) > ((a, b), b)$, $(b, (b, c)) > ((a, c), c)$ and H_3^o in any arbitrary way. For $n > 3$ the number of Hall elements grow very fast, but observe that in fact we are interested only in specific Hall elements. Namely, in H_n^a there will be only (but not all) elements of the form (v, w) , where $v \in H_k^a, w \in H_l^b, k, l < n$. This comes from the relations $[a, b] = 2a, [a, c] = -b, [b, c] = 2c$ satisfied in \mathfrak{A}_1 together with the assumptions (II) (in the definition of a Hall set) and (A) (the assumption on the ordering). Similarly, if $(v, w) \in H_n^b$, then $v \in H_k^a, w \in H_l^c, k, l < n$, and if $(v, w) \in H_n^c$, then $v \in H_k^b, w \in H_l^c, k, l < n$. Adding the assumption (III) of a Hall set, we reach the following lemma.

Lemma 4.1. *Assume that $H_1^a := \{a\}, H_1^b := \{b\}, H_1^c := \{c\}, H_1^o := \emptyset$. For $n > 1$ we recursively define ordered sets $H_n^d \subset M(A)$ by*

- (a) $H_n^a := \{(\dots(a, v_1) \dots, v_i) \in M(A) \mid v_i \in H^b, v_1 \geq \dots \geq v_i, |v_1| + \dots + |v_i| = n - 1\}$ with any total ordering.
- (b) $H_n^b := \{(a, v) \in M(A) \mid v \in H_{n-1}^c\}$ with ordering taken from H_{n-1}^c that is $(a, v) < (a, w)$ iff $v < w$;
- (c) (i) in case $n = 2k$ is even,

$$H_n^c := \bigcup_{i=1}^k \left\{ (v, w) \in M(A) \mid v \in H_i^b, w \in H_{n-i}^c \right\},$$

with "the lexicographical ordering" that is $(v, w) < (v', w')$ with $v, v' \in H_i^b$ and $w, w' \in H_{n-i}^c$ iff $v < v'$ or $(v = v' \text{ and } w < w')$;

- (ii) in case $n = 2k + 1$ is odd,

$$H_n^c := \bigcup_{i=1}^k \left\{ (v, w) \in M(A) \mid v \in H_i^b, w \in H_{n-i}^c \right\} \cup H_k^{odd},$$

where $H_k^{odd} := \{((a, v), w) \in M(A) \mid v, w \in H_k^c, v \geq w\}$. The ordering is like in case (i) with additional relations that $v < w$ for all $v \in H_k^{odd}, w \in H_n^c \setminus H_k^{odd}$, and $((a, v), w) < ((a, v'), w')$ with $v, v', w, w' \in H_k^c$ iff $v < v'$ or $(v = v' \text{ and } w < w')$.

- (d) $H_n^o := H_n^{aa} \cup H_n^{oa} \cup H_n^{bb} \cup H_n^{ob} \cup H_n^{cc} \cup H_n^{oc} \cup H_n^{oo}$ with any fixed ordering, where

$$H_n^{dd'} := \bigcup_{i=1}^{n-1} \left\{ (h', h'') \in M(A) \mid h' \in H_i^d, h'' \in H_{n-i}^{d'}, \text{ s.t. (II) and (III) are satisfied} \right\}.$$

Then $H^{\mathfrak{a}_1}$ defined by (I), (II) and (III) with the ordering \leq given by (A) and (B) is a Hall set.

Remark 4.2. The elements in H^o are not important for our purposes, and in fact we could consider any other elements which span the space $\text{span}\{P_h \mid h \in H^o\} \subset \text{Lie}(A)$ together with the elements in $H^a \cup H^b \cup H^c$ give a Hall set.

From (b) and (c) of the above lemma, we conclude that

Corollary 4.3. *In the set H_n^c there are only elements of three types, i.e.,*

- (a) $((a, v_1), v_2)$, where $v_1 \in H_{i-1}^c, v_2 \in H_{n-i}^c$, where $1 \leq i \leq (n+1)/2$, and $v_1 \geq v_2$;
- (b) (b, v) , where $v \in H_{n-1}^c$;
- (c) $c \in H_1^c$.

The above lemma is almost self explanatory, nevertheless we prove it.

Proof of Lemma 4.1. The condition (I) in the definition of a Hall set is obvious. Let us prove that the condition (II) is satisfied. We proceed by induction on the length n of the foliage of trees. The base assumption obviously comes from the condition (I). Assume $n > 1$. If $h \in H_n^a$, then it is of the form $h = (h', v)$, where $h' \in H_{n-|v|}^a$ and $v \in H_{|v|}^b$, so by (B) we have $h < v$. If $h \in H_n^b$, then $h = (a, v)$, where $v \in H_{n-1}^c$, so (B) implies $h < v$. If $h \in H_n^c$, then $h = (h', w)$, where $h' \in H_{n-|w|}^b$ and $w \in H_{|w|}^c$, so (A) implies $h < v$. For $h \in H_n^o$ there is nothing to prove.

Now let us check the condition (III). Assume $h = (h', h'') \in H^{\mathfrak{a}_1}$. Then $h \in H_n^a$ iff $h' \in H_{n-|h''|}^a$ and $h'' \in H_{|h''|}^b$, so by (B) $h' < h''$, and $h' = a$ or $h' = (x, y)$ with $x \in H_{|h'| - |y|}^a$ and $y \in H_{|y|}^b$ such that $y \geq h''$. Also $h \in H_n^b$ iff $h' = a$ and $h'' \in H_{n-1}^c$, so by (B) $a < h''$. Finally, $h \in H_n^c$ iff $h' \in H_{n-|h''|}^b$ and $h'' \in H_{|h''|}^c$, so by (B) $h' < h''$, and by the above corollary $h' = b$ or $h' = (x, y)$ with $x = a$, $y \in H_{|y|}^c$ and $y \geq h''$. For $h \in H_n^o$ once again there is nothing to prove. \square

Remark 4.4. The descriptions of the sets H_k^{odd} are not explicit. But observe that if we assume $H_k^c := \{v_1, \dots, v_q\}$ with the ordering $v_1 > \dots > v_q$, then it is easy to see that

$$H_k^{\text{odd}} = \bigcup_{i=1}^q \{((a, v_i), w) \in M(A) \mid w = v_i, v_{i+1}, \dots, v_q\}.$$

In Table 1 we show all the Hall words of type a, b and c of length less than 6. We abbreviate notation by omitting commas and the most exterior brackets.

H_n^d	$d = a$	$d = b$	$d = c$
$n = 1$	a	b	c
$n = 2$	(ab)	(ac)	(bc)
$n = 3$	$a(ac)$ $(ab)b$	$a(bc)$	$b(bc)$ $(ac)c$
$n = 4$	$a(a(bc))$ $(ab)(ac)$ $((ab)b)b$	$a(b(bc))$ $a((ac)c)$	$b(b(bc))$ $b((ac)c)$ $(ac)(bc)$
$n = 5$	$a(a(b(bc)))$ $a(a((ac)c))$ $(ab)(a(bc))$ $(a(ac))(ac)$ $((ab)b)(ac)$ $((ab)b)b$	$a(b(b(bc)))$ $a(b((ac)c))$ $a((ac)(bc))$	$b(b(b(bc)))$ $b(b((ac)c))$ $b((ac)(bc))$ $(ac)(b(bc))$ $(ac)((ac)c)$ $(a(bc))(bc)$
	\vdots	\vdots	\vdots

TABLE 1. Hall elements of type a, b and c of the length less than 6. The ordering is from up to down, and then from right to left.

5. PROOFS OF THEOREMS

In the previous section we constructed the Hall set $H^{\mathfrak{a}_1}$ suitable for our problem. Now we use it together with the Kawski-Sussmann solution (10). Namely, we assume that $'A = A = \{a, b, c\}$, that $Y_d = X_d$ for $d = a, b, c$, and that the Y_d 's satisfy (1). Using the Hall set $H^{\mathfrak{a}_1}$, we know by (10), that

$$(15) \quad x(t) = \overleftarrow{\prod}_{h \in H^{\mathfrak{a}_1}} \exp(\Phi_h(t) \cdot X_h)(x_0)$$

is a solution of (2). But H^{a_1} has special properties, which we now use. Namely, $H^{a_1} = H^o \cup H^a \cup H^b \cup H^c$, each H^d has the ordering of the type ω , and for $h_o \in H^o, h_a \in H^a, h_b \in H^b, h_c \in H^c$, we have $h_o < h_a < h_b < h_c$. Therefore,

$$\begin{aligned} \overleftarrow{\prod}_{h \in H^{a_1}} \exp(\Phi_h(t) \cdot X_h)(x_0) &= \overleftarrow{\prod}_{h \in H^c} \exp(\Phi_h(t) \cdot X_h) \overleftarrow{\prod}_{h \in H^b} \exp(\Phi_h(t) \cdot X_h) \\ &\quad \circ \overleftarrow{\prod}_{h \in H^a} \exp(\Phi_h(t) \cdot X_h) \overleftarrow{\prod}_{h \in H^o} \exp(\Phi_h(t) \cdot X_h)(x_0). \end{aligned}$$

Moreover, if $h \in H^d$ ($d = a, b, c$), then $X_h = \gamma_d(h) \cdot X_d$ with certain well defined functions $\gamma_d : H^d \rightarrow \mathbb{Z}$ (we give explicit expressions for γ_b and γ_c in Lemma 5.3), and if $h \in H^o$, then $X_h \equiv 0$. Thus it follows that

$$\begin{aligned} x(t) &= \overleftarrow{\prod}_{h \in H^{a_1}} \exp(\Phi_h(t) \cdot X_h)(x_0) = \exp\left(\sum_{h \in H^c} \Phi_h(t) \gamma_c(h) \cdot X_c\right) \\ &\quad \circ \exp\left(\sum_{h \in H^b} \Phi_h(t) \gamma_b(h) \cdot X_b\right) \exp\left(\sum_{h \in H^a} \Phi_h(t) \gamma_a(h) \cdot X_a\right)(x_0). \end{aligned}$$

Using the definition (13) of $\Phi_h(t)$ we get

$$\sum_{h \in H^d} \Phi_h(t) \gamma_d(h) = \Upsilon^t \left(\sum_{h \in H^d} \gamma_d(h) \xi_h \right)$$

(recall, that we use a natural abbreviation $\xi_h := \xi_{f(h)}$ for $h \in H^{a_1}$). In order to prove Theorem 1 we must show that

$$\sum_{h \in H^a} \gamma_a(h) \xi_h = a \exp_{\sqcup}(2S_{a_1}), \quad \sum_{h \in H^b} \gamma_b(h) \xi_h = S_{a_1}, \quad \sum_{h \in H^c} \gamma_c(h) \xi_h = \exp_{\sqcup}(2S_{a_1})c,$$

where $S_{a_1} \in \mathbb{R}\langle\langle A \rangle\rangle$ is the unique series satisfying $S_{a_1} = b - a \exp_{\sqcup}(2S_{a_1})c$. To do this, we first prove Theorem 1' (stated below) in which we simply change the series $S_{a_1} \in \mathbb{R}\langle\langle A \rangle\rangle$ from Theorem 1 into the main series S (defined in (5)), and then show that S is in fact the unique solution of the algebraic equation $S = b - a \exp_{\sqcup}(2S)c$, i.e., prove Proposition 2.4.

Theorem 1'. *Let $X_a, X_b, X_c \in \Gamma(M)$ be smooth tangent vector fields satisfying (1). Then (locally) the solution $x : [0, T] \rightarrow M$ of the differential equation (2) is of the form*

$$(4) \quad x(t) = \exp(\Xi_c(t)X_c) \exp(\Xi_b(t)X_b) \exp(\Xi_a(t)X_a)(x_0).$$

Here, $\Xi_a, \Xi_b, \Xi_c : [0, T] \rightarrow \mathbb{R}$ are given by $\Xi_d(t) := \Upsilon^t(S^d)$ (for $d = a, b, c$), where

$$S^a = a \exp_{\sqcup}(2S), \quad S^b = S, \quad S^c = \exp_{\sqcup}(2S)c,$$

and $S \in \mathbb{R}\langle\langle A \rangle\rangle$ is defined in (5). Additionally, we have

$$(16) \quad b - aS^c = S^b = b - S^a c.$$

5.1. Proof of Theorem 1'b with the first equality in (16). The following lemma demonstrates the connection between the mapping $s : W \rightarrow \mathbb{R}\langle A \rangle$, defined in section 2, and $\xi : H^{a_1} \rightarrow \mathbb{R}\langle A \rangle$, via the restriction of the correspondence between the set of words A^* and the dual basis in $\mathbb{R}\langle\langle A \rangle\rangle$.

Lemma 5.1. *Consider the Hall set H^{a_1} with its ordering \leq defined in section 4, and the set of words $W = \bigcup_{d=b,c} \bigcup_{n=1}^{\infty} W_n^d \subset A^*$ with ordering \leq defined in section 2. Then for $d = b, c$, and $n \in \mathbb{N}$, the Hall trees in H_n^d correspond with the Hall words in W_n^d , and the*

correspondence respects the orderings, i.e., for each $h \in M(A)$, $h \in H_n^d$ iff $f(h) \in W_n^d$, and $v < w$ iff $f(v) < f(w)$ for all $v, w \in H^b \cup H^c$. Moreover, for each $h \in W^b \cup W^c$,

$$\xi_h = s_h.$$

Remark 5.2. In other words, the above lemma states that the elements of W_n^d are the Hall words corresponding to the Hall trees in H_n^d , for $d = b, c$ and $n \in \mathbb{N}$.

Proof. The correspondence $h \in H_n^d \iff f(h) \in W_n^d$ (which respects the orderings) is as trivial as erasing the brackets for the Hall trees, which in fact the foliage mapping does. An inductive proof of the equality $\xi_h = s_h$ comes easily from Proposition 3.1, and the proved correspondence between elements in the Hall set and the Hall words, together with the definitions of these sets. \square

Lemma 5.3. *The functions $\gamma_b : H^b \rightarrow \mathbb{R}$ (defined in (14)) and $\alpha_b : W^b \rightarrow \mathbb{R}$ (defined in (6)) satisfy $\gamma_b(h) = \alpha_b(f(h))$ for all $h \in H^b$, i.e., $\gamma_b(b) = 1$ and $\gamma_b((a, v)) = -(-2)^{|f(v)|_a} \cdot 2^{|f(v)|_b}$ for each $(a, v) \in H^b$ ($v \in H^c$). Moreover, $\gamma_c(v) = -\gamma_b((a, v))$ for each $v \in H^c$.*

Proof. Recall that each $\gamma_d : H^d \rightarrow \mathbb{R}$ is defined by the formula $P_h = \gamma_d(h) \cdot d \bmod I(\mathfrak{a}_1)$. We first prove that $\gamma_c(v) = (-2)^{|f(v)|_a} \cdot 2^{|f(v)|_b}$ for each $v \in H^c$. We proceed by induction on the length of the foliage of v . For $v = c$ we obviously have $\gamma_c(c) = 1$. Assume the statement is satisfied for $v \in H_1^c \cup \dots \cup H_{n-1}^c$ and take $w \in H_n^c$. By Corollary 4.3 we must consider two cases.

Case I. $w = (a, v)$ where $v \in H_{n-1}^c$. Then

$$P_w = [b, P_v] = [b, \gamma_c(v) \cdot c] \bmod I(\mathfrak{a}_1) = 2\gamma_c(v) \cdot c \bmod I(\mathfrak{a}_1).$$

Therefore, using the inductive hypothesis, it follows that

$$\gamma_c(w) = 2 \cdot (-2)^{|f(v)|_a} \cdot 2^{|f(v)|_b} = (-2)^{|f(w)|_a} \cdot 2^{|f(w)|_b}.$$

Case II. $w = ((a, v_1), v_2)$ where $v_1 \in H_{i-1}^c$ and $v_2 \in H_{n-i}^c$. We compute

$$P_w = [[a, P_{v_1}], P_{v_2}] = -2\gamma_c(v_1)\gamma_c(v_2) \cdot c \bmod I(\mathfrak{a}_1).$$

By the inductive hypothesis we get

$$\gamma_c(w) = (-2) \cdot (-2)^{|f(v_1)|_a + |f(v_2)|_a} \cdot 2^{|f(v_1)|_b + |f(v_2)|_b} = (-2)^{|f(w)|_a} \cdot 2^{|f(w)|_b}.$$

Thus we have proved the formula for γ_c .

Now let us focus on $d = b$. Obviously $\gamma_b(b) = 1$. From Lemma 4.1, we know that an element $w \in H_n^b$ (for $n > 1$) is of the form $w = (a, v)$, where $v \in H_{n-1}^c$, so

$$P_w = [a, P_v] = -\gamma_c(v) \cdot b \bmod I(\mathfrak{a}_1).$$

Therefore $\gamma_b((a, v)) = \gamma_c(v)$, which together with the formula for γ_c , completes the proof. \square

From these two lemmas we conclude that

$$\sum_{h \in H^b} \gamma_b(h) \xi_h = \sum_{w \in W^b} \alpha_b(w) \xi_w = S,$$

which proves that $S^b = S$.

Additionally, from Lemma 4.1(b), Proposition 3.1, and Lemma 5.3, we get

$$\begin{aligned} \sum_{h \in H^b} \gamma_b(h) \xi_h &= \gamma_b(b) \xi_b + \sum_{v \in H^b} \gamma_b((a, v)) \xi_{av} \\ &= b + \sum_{v \in H^c} -\gamma_c(v) a \xi_v = b - a \sum_{v \in H^c} \gamma_c(v) \xi_v. \end{aligned}$$

Together with Lemma 5.1, the previous equation proves the first additional equality $b - a S^c = S^b$ of formula (16) in Theorem 1'.

5.2. Proof of Theorem 1'a and Proposition 2.6. In this section we calculate the expression S^a in the assertion of Theorem 1'.

From Lemma 4.1(a),

$$(17) \quad \begin{aligned} H^a &= \{a\} \cup \bigcup_{n=2}^{\infty} \left\{ (\cdots (a, v_1) \cdots, v_i) \mid v_i \in H^b, v_1 \geq \cdots \geq v_i, |v_1| + \cdots + |v_i| = n-1 \right\} \\ &= \left\{ (\cdots (a, v_1) \cdots, v_i) \in M(A) \mid i = 0, 1, 2, \dots, v_i \in H^b, v_1 \geq \cdots \geq v_i \right\}. \end{aligned}$$

Lemma 5.4. For $h = (\cdots (a, v_1) \cdots, v_i) \in H^a$ such that $v_i \in H^b$ and $v_1 \geq \cdots \geq v_i$, we have

$$\begin{aligned} (i) \quad \gamma_a(h) &= 2^i \cdot \gamma_b(v_1) \cdots \gamma_b(v_i) \\ (ii) \quad \xi_h &= a \xi_{v_1} \sqcup \cdots \sqcup \xi_{v_i}. \end{aligned}$$

The product \sqcup in $\mathbb{R}\langle\langle A \rangle\rangle$ is defined in analogy with the product defined in section 2. Namely, for $v_1 = \cdots = v_{i_1} > \cdots > w_1 = \cdots = w_{i_k}$, where $i_1, \dots, i_k \in \mathbb{N}$, we define

$$\xi_{v_1} \sqcup \cdots \sqcup \xi_{w_{i_k}} := \frac{1}{i_1! \cdots i_k!} \xi_{v_1}^{\sqcup i_1} \sqcup \cdots \sqcup \xi_{w_1}^{\sqcup i_k}.$$

Proof. (i) We proceed by induction on $i \in \mathbb{N}$. For $i = 0$, $\gamma_a(a) = 1$. Assume the assertion is satisfied for $i \leq j \in \mathbb{N}$, and take $h = (\cdots (a, v_1) \cdots, v_{j+1}) \in H^a$. By the definitions of γ_a and P , $P_h = [\cdots [a, v_1] \cdots, v_{j+1}] = \gamma_a(h) \cdot a \bmod I(\mathfrak{a}_1)$ and $P_{(a, v_1)} = [a, v_1] = \gamma_a((a, v_1)) \cdot a \bmod I(\mathfrak{a}_1)$. By the inductive hypothesis, the latter is equal to $2\gamma_b(v_1) \cdot a \bmod I(\mathfrak{a}_1)$, and it follows that

$$\begin{aligned} \gamma_a(h) \cdot a \bmod I(\mathfrak{a}_1) &= 2\gamma_b(v_1) \cdot [\cdots [a, v_2] \cdots, v_{j+1}] \bmod I(\mathfrak{a}_1) \\ &= 2\gamma_b(v_1) \cdot 2^j \cdot \gamma_b(v_2) \cdots \gamma_b(v_{j+1}) \cdot a \bmod I(\mathfrak{a}_1), \end{aligned}$$

where the inductive hypothesis is again used in the last line. Thus $\gamma_a(h) = 2^{j+1} \cdot \gamma_b(v_1) \cdots \gamma_b(v_{j+1})$ and we are done.

(ii) This is an immediate consequence of Proposition 3.1. \square

Using (17) together with the above lemma we get

$$\sum_{h \in H^a} \gamma_a(h) \xi_h = a \left(1 + \sum_{i=1}^{\infty} 2^i \sum_{\substack{v_1 \geq \cdots \geq v_i \\ v_i \in H^b}} \gamma_b(v_1) \cdots \gamma_b(v_i) \xi_{v_1} \sqcup \cdots \sqcup \xi_{v_i} \right).$$

Now, using Lemmas 5.1 and 5.3 with the notation $w_i = f(v_i)$, we get

$$\sum_{h \in H^a} \gamma_a(h) \xi_h = a \left(1 + \sum_{i=1}^{\infty} 2^i \sum_{\substack{w_1 \geq \cdots \geq w_i \\ w_i \in W^b}} \alpha_b(w_1) \cdots \alpha_b(w_i) s_{w_1} \sqcup \cdots \sqcup s_{w_i} \right),$$

where $\alpha_b(b) = 1$ and $\alpha_b(av) = -(-2)^{|v|_a} \cdot 2^{|v|_b}$ for $av \in W^b$ ($v \in W^c$). To end the proof of Theorem 1'(a), it remains to prove Proposition 2.6 which is the content of the following

calculation:

$$\begin{aligned}
\exp_{\sqcup}(2S) &= 1 + \sum_{i=1}^{\infty} \frac{2^i}{i!} \left(\sum_{w \in W^b} \alpha_b(w) s_w \right)^{\sqcup i} \\
&= 1 + \sum_{i=1}^{\infty} \frac{2^i}{i!} \left(\sum_{w_1 \in W^b} \alpha_b(w_1) s_{w_1} \right) \sqcup \cdots \sqcup \left(\sum_{w_i \in W^b} \alpha_b(w_i) s_{w_i} \right) \\
&= 1 + \sum_{i=1}^{\infty} 2^i \sum_{w_1, \dots, w_i \in W^b} \alpha_b(w_1) \cdots \alpha_b(w_i) \frac{1}{i!} s_{w_1} \sqcup \cdots \sqcup s_{w_i} \\
&= 1 + \sum_{i=1}^{\infty} 2^i \sum_{\substack{w_1 > \cdots > w_j \\ w_i \in W^b}} \sum_{\substack{(k_1, \dots, k_j) \in \mathbb{N}^j \\ k_1 + \dots + k_j = i}} (\alpha_b(w_1))^{k_1} \cdots (\alpha_b(w_j))^{k_j} \\
&\quad \times \frac{1}{k_1! \cdots k_j!} s_{w_1}^{\sqcup k_1} \sqcup \cdots \sqcup s_{w_j}^{\sqcup k_j}, \\
&= 1 + \sum_{i=1}^{\infty} 2^i \sum_{\substack{w_1 \geq \cdots \geq w_i \\ w_i \in W^b}} \alpha_b(w_1) \cdots \alpha_b(w_i) s_{w_1} \sqcup \cdots \sqcup s_{w_i}.
\end{aligned}$$

Remark 5.5. Note that the above calculation is in fact very similar to the one Kawski and Sussmann gave proving their theorem [KS97].

5.3. End of proof of Theorem 1'. In this section we calculate the expression S^c and the second additional equality in the assertion of Theorem 1'. As a byproduct we also get Theorem 2.

At this point we know that (locally) the solution of the system

$$\begin{aligned}
(2) \quad \dot{x}(t) &= u_c(t)X_c + u_b(t)X_b + u_a(t)X_a, \\
x(0) &= x_0 \in M.
\end{aligned}$$

can be written in the form

$$(18) \quad x(t) = \exp \left(\tilde{\Xi}_c(t)X_c \right) \exp \left(\Xi_b(t)X_b \right) \exp \left(\Xi_a(t)X_a \right) (x_0),$$

where

$$\tilde{\Xi}_c(t) = \Upsilon^t \left(\sum_{h \in H^c} \gamma_c(h) \xi_h \right),$$

and $\Xi_a(t) = \Upsilon^t(a \exp_{\sqcup}(2S))$, $\Xi_b(t) = \Upsilon^t(S)$. Differentiating (18) with respect to t gives the following equation:

$$\begin{aligned}
\dot{x} &= (-\Xi_a^2 e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c - 2\Xi_a \cdot \dot{\Xi}_b + \dot{\Xi}_a)X_a \\
&\quad + (\Xi_a e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c + \dot{\Xi}_b)X_b \\
&\quad + (e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c)X_c.
\end{aligned}$$

Comparing this equation with (2), we see that

$$\begin{aligned}
(19) \quad u_a &= -\Xi_a^2 e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c - 2\Xi_a \cdot \dot{\Xi}_b + \dot{\Xi}_a, \\
u_b &= \Xi_a e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c + \dot{\Xi}_b, \\
u_c &= e^{-2\Xi_b} \cdot \dot{\tilde{\Xi}}_c.
\end{aligned}$$

Considering this as a linear system of equations in variables $\Xi_a, \Xi_b, \tilde{\Xi}_c$, it is easy to conclude that the first equation in (19) can be rewritten in the form

$$(20) \quad \dot{\Xi}_a = u_a + 2u_b \cdot \Xi_a - u_c \cdot \Xi_a^2.$$

Since $\Xi_a(0) = 0$, we have proved the following theorem.

Theorem 2'. *For fixed measurable functions $u_a, u_b, u_c : [0, T] \rightarrow \mathbb{R}$, the function $\Xi_a : [0, T] \rightarrow \mathbb{R}$, defined in Theorem 1 by $\Xi_a(t) = \Upsilon^t(a \exp_{\sqcup}(2S))$, is (locally) the solution of the Riccati equation $\dot{\Xi}_a(t) = u_a(t) + 2u_b(t) \cdot \Xi_a(t) - u_c(t) \cdot \Xi_a^2(t)$, $\Xi_a(0) = 0$.*

Obviously, combining the second and the third equation in (19) we get

$$(21) \quad \dot{\Xi}_b = u_b - u_c \cdot \Xi_a, \quad \dot{\tilde{\Xi}}_c = u_c \cdot e^{2\Xi_b}.$$

Integrating the first equation and using the definition (3) of Υ^t we conclude

$$\Xi_b(t) = \Upsilon^t(b - S^a c).$$

Combining it with the defining formula $\Xi_b(t) = \Upsilon^t(S^b)$, we obtain $S^b = b - S^a c$ which is the second of the additional equality of Theorem 1'. Finally, from the second equation in (21), we obtain that $S^c = \exp_{\sqcup}(2S) c$. To see this, we use (12) to get that

$$e^{2\Xi_b} = 1 + \sum_{k=1}^{\infty} \frac{2^k}{k!} (\Upsilon^t(S))^k = 1 + \sum_{k=1}^{\infty} \frac{2^k}{k!} \Upsilon^t(S^{\sqcup k}) = \Upsilon^t(\exp_{\sqcup}(2S)).$$

Therefore, integrating the second equality in (21), we conclude that

$$\tilde{\Xi}_c(t) = \int_0^t u_c(t_1) \cdot \Upsilon^{t_1}(\exp_{\sqcup}(2S)) dt_1 = \Upsilon^{t_1}(\exp_{\sqcup}(2S)c),$$

so $S^c = \exp_{\sqcup}(2S) c$.

Remark 5.6. The system of three equations given by (20) and (21) was actually considered by Redheffer [Red56] (see also [Red57]), who demonstrated a connection between this system and a non-autonomous linear system on \mathbb{C}^2 with $\mathfrak{sl}(2)$ -matrix valued functions of time. Our results reproduce this phenomena and go further in producing a solution to the system, i.e., Theorem 2.

5.4. Proof of Proposition 2.4 and Theorems 1 and 2. From Theorem 1' we easily conclude that

$$S = S^b = b - S^a c = b - a \exp_{\sqcup}(2S) c,$$

so $S \in \mathbb{R}\langle\langle A \rangle\rangle$ is a solution of the equation given in the proposition. We prove uniqueness of this solution. Let $S = \sum_{n=1}^{\infty} S_n$, where $S_n \in \mathbb{R}\langle\langle A \rangle\rangle$ is homogeneous of degree n , be a solution of the above equation. Then $S_1 = b$, $S_2 = -ac$, and for $n > 2$

$$S_n = -a \left(2S_{n-2} + \frac{2^2}{2!} \sum_{k+l=n-2} S_k \sqcup S_l + \frac{2^3}{3!} \sum_{k+l+m=n-2} S_k \sqcup S_l \sqcup S_m + \dots \right) c,$$

which gives us a recursive definition of the homogeneous summands in S . Therefore S is defined uniquely.

Once again we emphasize that Theorem 1' together with Proposition 2.4 imply Theorem 1. It is also obvious that Theorem 2' (proved in section 5.3) is now equivalent to Theorem 2.

5.5. Proof of Proposition 2.5. From Proposition 2.4 we know that the main series $S \in \mathbb{R}\langle\langle A \rangle\rangle$ is the unique solution of $S = b - a \exp_{\sqcup}(2S)c$. Since $(\cdot)^*$ is an algebra antihomomorphism, and \sqcup is commutative, we obtain $S^* = b - c \exp_{\sqcup}(2S^*)a$. So using the definition of F , we conclude that $F^*(S) = b - a \exp_{\sqcup}(2F^*(S))c$. It follows that $F^*(S)$ satisfies the equation from Proposition 2.4 for which S is the unique solution, so $F^*(S) = S$. Now the proposition follows easily from the equations established in Theorem 1.

6. CONCLUDING REMARKS

In this article we presented the explicit solution of a general non-linear control-affine system of \mathfrak{a}_1 -type. The solution was given in terms of the composition of three flows, with additional non-linear dependence on time, written with the use of certain well defined non-commutative power series S on three letters.

Let us underline, that in order to formulate the results, we did not use the Hall basis structure, but only the set of Hall words. In other words, the solution of our problem does not depend on the edges of the trees. This fact makes the results easy to formulate, and it is also consistent with intuition, i.e., the solution should not depend on a chosen basis.

The solution given for the simplest simple algebra gives hope for similar solutions for other simple algebras. The first step is to find expressions for four rank-two simple Lie algebras of type $\mathfrak{a}_2, \mathfrak{b}_2, \mathfrak{c}_2$ and \mathfrak{g}_2 . This will be a topic of the author's future research.

Finally, let us mention that the approach given in this article can not be easily generalized to the case of the special unitary algebra $\mathfrak{su}(2)$ (and $\mathfrak{su}(n)$), which would be helpful in studies of quantum (control) systems. The problem comes from the fact that generators of this Lie algebra (i.e., the Pauli matrices) are in a cyclic relation. It will definitely be interesting to overcome this problem.

ACKNOWLEDGEMENTS

The author was partially supported by the Polish Ministry of Research and Higher Education grant N201 607540, 2011-2014.

REFERENCES

- [AS04] Andrei A. Agrachev and Yuri L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, vol. 87, Springer-Verlag, Berlin, 2004, Control Theory and Optimization, II. MR 2062547 (2005b:93002)
- [Cd11] J. F. Cariñena and J. de Lucas, *Integrability of Lie systems through Riccati equations*, Journal of Nonlinear Mathematical Physics **18** (2011), no. 1, 29–54.
- [CdL09] José F. Cariñena and Javier de Lucas, *Applications of Lie systems in dissipative Milne-Pinney equations*, Int. J. Geom. Methods Mod. Phys. **6** (2009), no. 4, 683–699. MR 2541945 (2010f:34007)
- [CdLR08] José F. Cariñena, Javier de Lucas, and Manuel F. Rañada, *Recent applications of the theory of Lie systems in Ermakov systems*, SIGMA Symmetry Integrability Geom. Methods Appl. **4** (2008), Paper 031, 18. MR 2393296 (2009a:34013)
- [CdLR10] ———, *Lie systems and integrability conditions for t -dependent frequency harmonic oscillators*, Int. J. Geom. Methods Mod. Phys. **7** (2010), no. 2, 289–310. MR 2646785
- [Che54] Kuo-Tsai Chen, *Iterated integrals and exponential homomorphisms*, Proc. London Math. Soc. (3) **4** (1954), 502–512. MR 0073174 (17,394g)
- [Che57] ———, *Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula*, Ann. of Math. (2) **65** (1957), 163–178. MR 0085251 (19,12a)
- [Che68] ———, *Algebraic paths*, J. Algebra **10** (1968), 8–36. MR 0229630 (37 #5204)
- [CLR07a] José F. Cariñena, Javier de Lucas, and Arturo Ramos, *A geometric approach to integrability conditions for Riccati equations*, Electron. J. Differential Equations (2007), No. 122, 14 pp. (electronic). MR 2349950 (2008i:34012)
- [CLR07b] ———, *A geometric approach to integrability conditions for Riccati equations*, Electron. J. Differential Equations (2007), No. 122, 14 pp. (electronic). MR 2349950 (2008i:34012)
- [Fli81] Michel Fliess, *Fonctionnelles causales non linéaires et indéterminées non commutatives*, Bull. Soc. Math. France **109** (1981), no. 1, 3–40. MR 613847 (82h:93037)

- [Kaw02] Matthias Kawski, *The combinatorics of nonlinear controllability and noncommuting flows*, Mathematical control theory, Part 1, 2 (Trieste, 2001), ICTP Lect. Notes, VIII, Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002, pp. 223–311 (electronic). MR 1972790
- [KS97] Matthias Kawski and Héctor J. Sussmann, *Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory*, Operators, systems, and linear algebra (Kaiserslautern, 1997), European Consort. Math. Indust., Teubner, Stuttgart, 1997, pp. 111–128. MR 1603219 (99a:93021)
- [MR89] Guy Melançon and Christophe Reutenauer, *Lyndon words, free algebras and shuffles*, Canad. J. Math. **41** (1989), no. 4, 577–591. MR 1012617 (90j:20074)
- [Red56] R. M. Redheffer, *On solutions of Riccati's equation as functions of the initial values*, J. Rational Mech. Anal. **5** (1956), 835–848. MR 0088664 (19,558c)
- [Red57] ———, *The Riccati equation: Initial values and inequalities*, Math. Ann. **133** (1957), 235–250. MR 0089336 (19,655a)
- [Reu93] Christophe Reutenauer, *Free Lie algebras*, London Mathematical Society Monographs. New Series, vol. 7, The Clarendon Press Oxford University Press, New York, 1993, Oxford Science Publications. MR 1231799 (94j:17002)
- [Shi62] A. I. Shirshov, *On the bases of a free lie algebra*, Algebra Logika **1** (1962), no. 1, 14–19.
- [Sus86] H. J. Sussmann, *A product expansion for the Chen series*, Theory and applications of nonlinear control systems (Stockholm, 1985), North-Holland, Amsterdam, 1986, pp. 323–335. MR 935387 (89c:93030)
- [Vie78] Gérard Viennot, *Algèbres de Lie libres et monoïdes libres*, Lecture Notes in Mathematics, vol. 691, Springer, Berlin, 1978, Bases des algèbres de Lie libres et factorisations des monoïdes libres. MR 516004 (80i:17016)
- [WN64] J. Wei and E. Norman, *On global representations of the solutions of linear differential equations as a product of exponentials*, Proc. Amer. Math. Soc. **15** (1964), 327–334. MR 0160009 (28 #3223)

INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES; ŚNIADECKICH 8, 00-956 WARSZAWA, POLAND; ☎+48 22 5544524.

E-mail address: G.Pietrzkowski@impan.pl